

Sujet : Détection d'anomalies dans des données de graphe peu étiquetées.

Description:

La détection de nœuds anormaux dans un graphe présente d'importants défis en raison de la rareté des anomalies et du coût élevé associé à l'étiquetage de ce genre de données [1,2]. Ces difficultés engendrent des problèmes centrés sur les données, tels que des distributions extrêmement déséquilibrées et un apprentissage complexe, ce qui entrave fortement les méthodes d'apprentissage automatique et d'apprentissage [3,4]. Il existe, dans la littérature, quelques méthodes dédiées à ce problème [5,6] et le but de ce stage est de les étudier et les comparer sur différents niveaux de pénurie d'étiquettes. Les objectifs de ce stage sont :

- Etudier les méthodes de détection d'anomalies sur les données de types graphes avec peu d'étiquettes.
- Comparer les méthodes existantes sur des datasets de référence
- Proposer des améliorations

Références :

[1] Wang Y, Yang Z, Che X (2025) A Hierarchical Mixture-Of-Experts Framework for Few Labeled Node Classification. *Neural networks* 10.1016/j.neunet.2025.107285188(107285)

[2] Zhihui Hu, Yao Fu, Hong Zhao, Xiaoyu Cai, Weihao Jiang, and Shiliang Pu. 2023. Liberate Pseudo Labels from Over-Dependence: Label Information Migration on Sparsely Labeled Graphs. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM '23)*. Association for Computing Machinery, New York, NY, USA, 833–842. <https://doi.org/10.1145/3583780.3614954>

[3] Eunji Koh, Young Jae Lee, Seoung Bum Kim, GraphixMatch: Improving semi-supervised learning for graph classification with FixMatch, *Neurocomputing*, Volume 607, 2024, 128356, <https://doi.org/10.1016/j.neucom.2024.128356>.

[4] Sahil Manchanda, Shubham Gupta, Sayan Ranu, Srikanta J. Bedathur. Generative modeling of labeled graphs under data scarcity. *Proceedings of the Second Learning on Graphs Conference*, PMLR 231:32:1-32:18, 2024.

[5] Aoge Chen, Jianshe Wu, Hongtao Zhang, FIAD: Graph anomaly detection framework based feature injection, *Expert Systems with Applications*, Volume 259, 2025, 125216, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2024.125216>.

[6] Xiaoxiao Ma, Ruikun Li, Fanzhen Liu, Kaize Ding, Jian Yang, and Jia Wu. 2024. Graph Anomaly Detection with Few Labels: A Data-Centric Approach. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)*. Association for Computing Machinery, New York, NY, USA, 2153–2164. <https://doi.org/10.1145/3637528.3671929>